ON CERTAIN FINITE-DIFFERENCE METHODS FOR FLUID DYNAMICS

AKIRA KASAHARA
National Center for Atmospheric Research, Boulder, Colo.

ABSTRACT

Two finite-difference methods for geophysical fluid problems are described, and stability conditions of these schemes are discussed. These two schemes are formulated based upon a similar procedure given by Lax and Wendroff in order to obtain a second-order accuracy in finite-difference equations. However, the two schemes show remarkable differences in their computational stability. One scheme is stable, as one might expect, under the usual stability conditions of Courant-Friedrichs-Lewy and Lax-Wendroff. However, the other scheme is conditionally stable only if the flow is supercritical (supersonic in the case of gas dynamics) and unconditionally unstable if the flow is subcritical (subsonic).

1. INTRODUCTION

In geophysical fluid problems, one must often solve numerically the partial differential equations that govern the one-dimensional motion of a homogeneous incompressible fluid (e.g., Stoker [8]),

$$\frac{\partial u}{\partial t}=-u\frac{\partial u}{\partial x}-g\frac{\partial h}{\partial x},$$

$$\frac{\partial h}{\partial t}=-u\frac{\partial h}{\partial x}-h\frac{\partial u}{\partial x}, \tag{1}$$

where t and x denote the time and the space variable. $u(=u(x,t))$ and $h(=h(x,t))$ represent the speed and the depth of the fluid, and g stands for the acceleration due to gravity. For this first-order system, the characteristic directions $\tau=dx{:}dt$ are defined by

$$\begin{vmatrix} u-\tau & g \\ h & u-\tau \end{vmatrix}=0,$$

hence $\tau=u\pm\sqrt{gh}$. The system is hyperbolic since there are two distinct real roots τ (Courant and Hilbert [2]).

Richtmyer [7] made a survey of difference methods which are applicable to a hyperbolic system such as (1). In this paper, we shall describe two additional finite-difference methods of second-order accuracy which are formulated based upon a similar procedure given by Lax and Wendroff [4]. Our main concern will be to point out remarkable differences in the computational stability of the two schemes in spite of close similarity. In this discussion, stability will be taken to mean the stability of the corresponding linearized system with constant coefficients (Richtmyer [5]).

The linearized equations of (1) may be written as

$$\frac{\partial \hat{u}}{\partial t}=-U\frac{\partial \hat{u}}{\partial x}-C\frac{\partial \hat{h}}{\partial x},$$

$$\frac{\partial \hat{h}}{\partial t}=-U\frac{\partial \hat{h}}{\partial x}-C\frac{\partial \hat{u}}{\partial x}, \tag{2}$$

where $C=\sqrt{gH}$, H denotes a constant depth of the fluid, $\hat{u}=u/C$ and $\hat{h}=h/H$. Here U denotes a constant speed of the fluid. From this point on, we shall omit writing circumflex symbols for dimensionless variables $\hat{u}$ and $\hat{h}$ whenever references are made to (2).

In order to write down difference equations, we shall use a rectangular net in the x-t plane, with spacings Δx and Δt. We abbreviate any function $f(x,t)$ of $x=l\Delta x$ and $t=m\Delta t$ as f_l^m or $[f]_l^m$ where l and m can be either an integer or half an odd integer.

2. FINITE-DIFFERENCE METHOD I

In this scheme, there will occur values of u at integer space points and half-odd-integer times, and values of h at half-odd-integer space points and integer times as illustrated in figure 1. The following scheme was suggested by Richtmyer [6], but its stability condition was not discussed. The stability of a similar method is investigated by Fischer [3], but the analysis is limited to long-wave Fourier components.

The difference form of (2) may be written as

$$u_j^{n+1/2}=u_j^{n-1/2}-U\left\{\frac{\partial u}{\partial x}\right\}_j^n\Delta t-C\left(\frac{\partial h}{\partial x}\right)_j^n t\Delta,$$

$$h_{j+1/2}^{n+1}=h_{j+1/2}^{n}-U\left\{\frac{\partial h}{\partial x}\right\}_{j+1/2}^{n+1/2}\Delta t-C\left(\frac{\partial u}{\partial x}\right)_{j+1/2}^{n+1/2}\Delta t, \tag{3}$$*

*In the difference equations described in this note, the notations

$$\left(\frac{\partial f}{\partial x}\right)_l^m,\ \left\{\frac{\partial f}{\partial x}\right\}_l^m,\ \left[\frac{\partial f}{\partial x}\right]_l^m,\text{ and }\left\langle\frac{\partial f}{\partial x}\right\rangle_l^m$$

all have different meanings as defined respectively by (4), (5), in (5), and by (7).

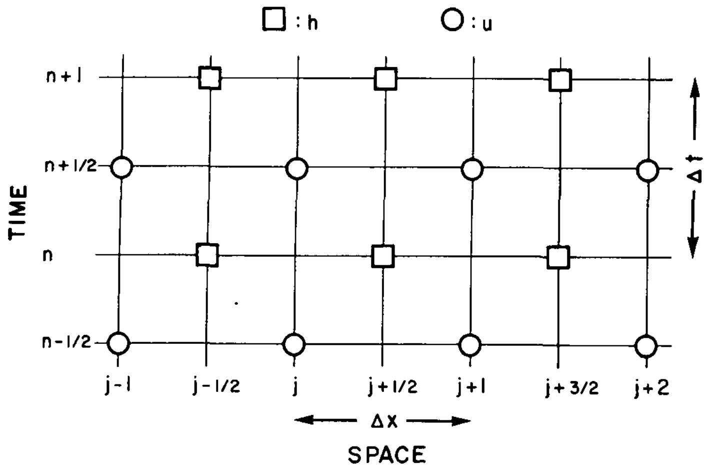

FIGURE 1.—Lattice structure for method I: n and j are integers.

where n and j are integers, and $(\partial f/\partial x)_l^m$ denotes the evaluation of $\partial f/\partial x$ at the time level m and the space point l. For second-order accuracy in Δx, we use the centered approximation

$$(\partial f/\partial x)_l^m = (f_{l+1/2}^m - f_{l-1/2}^m)/\Delta x \tag{4}$$

with Δx as the difference interval. (This can be done, because the space point l falls in the middle of the two adjacent points which carry values of f.)

In order to evaluate $\partial u/\partial x$ at integer time levels (when only h's are available) and $\partial h/\partial x$ at half-odd-integer time levels (when only u's are available), we expand these derivatives into Taylor series in time and retain enough terms to ensure that (3) has second-order accuracy in Δ, as was done by Lax and Wendroff [4].

The results are

$$\left\{\frac{\partial u}{\partial x}\right\}_j^n = \left[\frac{\partial u}{\partial x}\right]_j^{n-1/2} + \frac{\Delta t}{2}\frac{\partial}{\partial t}\left[\frac{\partial u}{\partial x}\right]_j^{n-1/2} = \left\langle\frac{\partial u}{\partial x}\right\rangle_j^{n-1/2} - \frac{\Delta t}{2}\left[U\left(\frac{\partial^2 u}{\partial x^2}\right)_j^{n-1/2} + C\widetilde{\left(\frac{\partial^2 h}{\partial x^2}\right)_j^n}\right] + O(\Delta^3) \tag{5}$$

and

$$\left\{\frac{\partial h}{\partial x}\right\}_{j+1/2}^{n+1/2} = \left[\frac{\partial h}{\partial x}\right]_{j+1/2}^n + \frac{\Delta t}{2}\frac{\partial}{\partial t}\left[\frac{\partial h}{\partial x}\right]_{j+1/2}^n = \left\langle\frac{\partial h}{\partial x}\right\rangle_{j+1/2}^n - \frac{\Delta t}{2}\left[U\left(\frac{\partial^2 h}{\partial x^2}\right)_{j+1/2}^n + C\widetilde{\left(\frac{\partial^2 u}{\partial x^2}\right)_{j+1/2}^{n+1/2}}\right] + O(\Delta^3), \tag{6}$$

where $\langle\partial f/\partial x\rangle_l^m$ denotes the evaluation of $\partial f/\partial x$ at the time level m and the space point l. In contrast to the formula (4), we use $2\Delta x$ as the difference interval and approximate

$$\left\langle\frac{\partial f}{\partial x}\right\rangle_l^m = \frac{f_{l+1}^m - f_{l-1}^m}{2\Delta x}. \tag{7}$$

(This is done because the space point l coincides with one of the points at which values of f appear.) Note that (2) is used to eliminate the time dependent terms in (5) and (6). The evaluation of the second term in the brackets of (5) and of (6) is not made at the same time level as for the first term in the brackets, but this approximation causes an error of only third order in Δ which can be neglected in the scheme of second-order accuracy. Since h's are not available at integer space points and u's are not at half-odd-integer space points, the second term in the brackets of (5) and (6) may be evaluated with the following space averaging,

$$\widetilde{\left(\frac{\partial^2 f}{\partial x^2}\right)_l^m} = \frac{1}{2}\left[\left(\frac{\partial^2 f}{\partial x^2}\right)_{l+1/2}^m + \left(\frac{\partial^2 f}{\partial x^2}\right)_{l-1/2}^m\right], \tag{8}$$

where

$$\left(\frac{\partial^2 f}{\partial x^2}\right)_l^m = (f_{l+1}^m - 2f_l^m + f_{l-1}^m)/(\Delta x)^2. \tag{9}$$

Let us substitute in (3) typical Fourier terms

$$u_j^{n+1/2} = u^{n+1/2} e^{ik(j\Delta x)}$$

$$h_{j+1/2}^n = h^n e^{ik\{(j+1/2)\Delta x\}}.$$

With the aid of formulas (4)–(9) and calling

$$Q = i\sin\theta + \alpha(1-\cos\theta),$$

$$P = (2i + \alpha\sin\theta)\sin\frac{\theta}{2},$$

$$\alpha = U\frac{\Delta t}{\Delta x},\ \beta = C\frac{\Delta t}{\Delta x},\ \theta = k\Delta x,$$

$$a = 1 - \alpha Q,\ d = -\beta P, \tag{10}$$

we obtain from (3),

$$u^{n+1/2} = au^{n-1/2} + dh^n, \tag{11}$$

$$h^{n+1} = du^{n+1/2} + ah^n. \tag{12}$$

By eliminating $u^{n+1/2}$ from (12) with the use of (11), we have

$$\begin{pmatrix} u^{n+1/2} \\ h^{n+1} \end{pmatrix} = G\begin{pmatrix} u^{n-1/2} \\ h^n \end{pmatrix}$$

where

$$G \equiv \begin{pmatrix} a & d \\ ad & a+d^2 \end{pmatrix} \tag{13}$$

is the amplification matrix. The von Neumann stability condition requires that the eigenvalues of the amplification matrix should not exceed unity in absolute value for physically stable systems (e.g., Richtmyer [5]). The eigenvalues of (13) are the roots of

$$\lambda^2 - (2a + d^2)\lambda + a^2 = 0. \tag{14}$$

Although this is only a quadratic equation, the fact that a and d are complex makes it difficult to see the conditions for which $|\lambda| \leq 1$. We, therefore, will discuss special cases first.

Case I.1 in which $C=0$.

In case the gravity-wave speed C vanishes, we have $d=0$ and the two equations (11) and (12) are uncoupled. Equation (14) reduces to $(\lambda-a)^2=0$ and

$$|a|^2=\{1-\alpha^2(1-\cos\theta)\}^2+\alpha^2\sin^2\theta,$$
$$=1-\alpha^2(1-\alpha^2)(1-\cos\theta)^2.$$

We wish $|\lambda|$ to be equal or less than unity; hence we must have $|\alpha|\leq 1$ or

$$\left|U\frac{\Delta t}{\Delta x}\right|\leq 1 \tag{15}$$

for stability as discussed by Lax and Wendroff [4].

Case I.2 in which $U=0$.

In case the advective flow speed U vanishes, equation (14) reduces to

$$\lambda^2-2\left(1-2\beta^2\sin^2\frac{\theta}{2}\right)\lambda+1=0.$$

To keep the roots of this equation in or on the unit circle, we must have $\left|\beta\sin\frac{\theta}{2}\right|\leq 1$ or

$$C\frac{\Delta t}{\Delta x}\leq 1 \tag{16}$$

which is the well-known stability condition by Courant, Friedrichs, and Lewy [1].

Case I.3 in which $U\neq C\neq 0$.

In this general case, the roots of (14) were computed numerically for various values of U, C, and θ defined in (10). In figure 2, the magnitude of the largest root is plotted against $|\alpha|$ as the abscissa and β as the ordinate. This largest root is found for $\theta=\pi$, namely $k\Delta x=\pi$. Since k is the wave number defined by $k\equiv 2\pi/L$, where L is the wavelength, the case of $\theta=\pi$ corresponds to that of $L=2\Delta x$, the shortest wavelength which the grid can resolve. In this case, we have from (10) that

$$Q=2\alpha,\ P=2i,$$
$$a=1-2\alpha^2,\ d=-2\beta i.$$

Then, equation (14) reduces to

$$\lambda^2-2\{1-2(\alpha^2+\beta^2)\}\lambda+(1-2\alpha^2)^2=0.$$

The roots of this equation are

$$\lambda=1-2(\alpha^2+\beta^2)\pm 2\beta\sqrt{2\alpha^2+\beta^2-1}.$$

In figure 2, the curve $1=2\alpha^2+\beta^2$ is shown by the chain line running through the trough of the contours. For $2\alpha^2+\beta^2<1$ on the left-hand side of the chain line, the two roots become complex conjugate and their magnitudes are both equal to $|1-2\alpha^2|$ which does not depend upon the parameter β. For $2\alpha^2+\beta^2>1$ on the right-hand side of the chain line in figure 2, it can be shown that the two

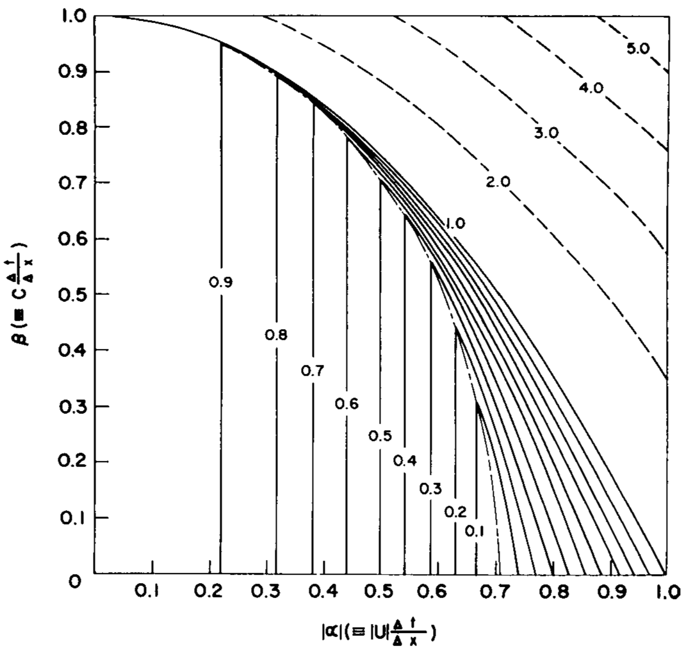

FIGURE 2.—Magnitude of the largest root ($\theta=\pi$) of equation (14), plotted against $|\alpha|$ and β (method I).

roots are real and negative; one of the roots becomes equal to -1 for $\alpha^2+\beta=1$ and the roots of (14) must all lie in or on the unit circle for

$$\left(U\frac{\Delta t}{\Delta x}\right)^2+C\frac{\Delta t}{\Delta x}\leq 1. \tag{17}$$

One might say that the above stability condition is a reasonable one and in fact one can "guess" intuitively this kind of result from the stability conditions (15) and (16) of the two special cases. However, the stability analysis of the next method will demonstrate an example that such a "guess" does not necessarily work.

3. FINITE-DIFFERENCE METHOD II

We will now modify the method I in the following manner. In this scheme, there will appear values of u at half-odd-integer times and integer space points (just as in method I). However, values of h appear also at integer space points but only at integer times (therefore half-odd-integer space points are removed) as illustrated in figure 3. The difference form of (2) may now be written

$$u_j^{n+1/2}=u_j^{n-1/2}-U\left\{\frac{\partial u}{\partial x}\right\}_j^n\Delta t-C\left\langle\frac{\partial h}{\partial x}\right\rangle_j^n\Delta t,$$
$$h_j^{n+1}=h_j^n-U\left\{\frac{\partial h}{\partial x}\right\}_j^{n+1/2}\Delta t-C\left\langle\frac{\partial u}{\partial x}\right\rangle_j^{n+1/2}\Delta t. \tag{18}$$

Here, we have used the same notations as introduced in method I.

As discussed in connection with (5) and (6), we expand similarly

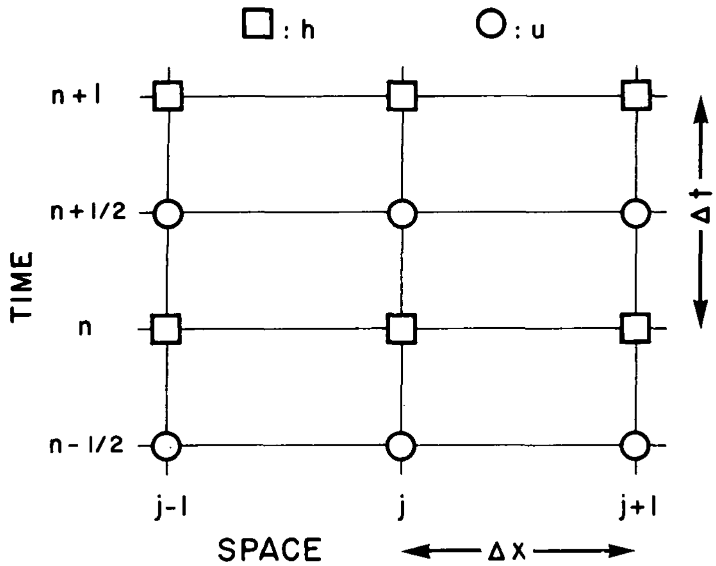

FIGURE 3.—Lattice structure for method II: n and j are integers.

$$\left\{\frac{\partial u}{\partial x}\right\}_j^n=\left\langle\frac{\partial u}{\partial x}\right\rangle_j^{n-1/2}-\frac{\Delta t}{2}\left[U\left(\frac{\partial^2 u}{\partial x^2}\right)_j^{n-1/2}+C\left(\frac{\partial^2 h}{\partial x^2}\right)_j^n\right]+O(\Delta^3), \quad (19)$$

$$\left\{\frac{\partial h}{\partial x}\right\}_j^{n+1/2}=\left\langle\frac{\partial h}{\partial x}\right\rangle_j^n-\frac{\Delta t}{2}\left[U\left(\frac{\partial^2 h}{\partial x^2}\right)_j^n+C\left(\frac{\partial^2 u}{\partial x^2}\right)_j^{n+1/2}\right]+O(\Delta^3). \quad (20)$$

Note that the second term in the brackets of (19) and (20) can be computed at j-points directly. The first and second derivatives are evaluated with the centered difference formulas (4), (7), and (9).

By introducing a typical Fourier term

$$f_j^m=f^m e^{ik(j\Delta x)}$$

into (18) and taking into account (19) and (20) with formulas (4), (7), and (9), and calling

$$Q=i\sin\theta+\alpha(1-\cos\theta),$$
$$\alpha=U\frac{\Delta t}{\Delta x},\ \beta=C\frac{\Delta t}{\Delta x},\ \theta=k\Delta x,$$
$$a=1-\alpha Q,\ b=-\beta Q, \quad (21)$$

we obtain

$$\begin{pmatrix}u^{n+1/2}\\ h^{n+1}\end{pmatrix}=G\begin{pmatrix}u^{n-1/2}\\ h^n\end{pmatrix}$$

where

$$G\equiv\begin{pmatrix}a & b\\ ab & a+b^2\end{pmatrix}.$$

The eigenvalues of G are the roots of the equation

$$\lambda^2-(2a+b^2)\lambda+a^2=0. \quad (22)$$

Case II.1 in which $C=0$.

The stability condition in this case is exactly the same as that of Case I.1, namely

$$\left|U\frac{\Delta t}{\Delta x}\right|\leq 1. \quad (23)$$

Case II.2 in which $U=0$.

In this case, equation (22) reduces to

$$\lambda^2-(2-\beta^2\sin^2\theta)\lambda+1=0. \quad (24)$$

To keep the roots of (24) in or on the unit circle, we must have $|\beta\sin\theta|\leq 2$ or

$$C\frac{\Delta t}{(2\Delta x)}\leq 1 \quad (25)$$

which corresponds to the $C-F-L$ condition (16) of Case I.2.

Case II.3 in which $|U|=C\neq 0$.

In this case, we have from (21) that $a=1-\alpha Q$ and $b=\mp\alpha Q$. Equation (22) reduces to

$$\lambda^2-(1+a^2)\lambda+a^2=0. \quad (26)$$

One of the roots of (26) is a^2 and the other is unity! Therefore the stability in this case is determined only from the condition that $|a|\leq 1$ which leads to the same condition as (15) discussed in Case I.1.

Case II.4 in which $U\neq C\neq 0$.

In this general case, the roots of (22) were calculated numerically for various values of U, C, and θ defined in (21). In figure 4, the magnitude of the largest root of (22) is plotted against $|\alpha|$ as the abscissa and β as the ordinate. This largest root is obtained for $\theta=\pi$. In this case, we have from (21) that

$$Q=2\alpha,\ a=1-2\alpha^2,\ b=-2\beta\alpha.$$

Equation (22) then reduces to

$$\lambda^2-2(1-2\alpha^2+2\alpha^2\beta^2)\lambda+(1-2\alpha^2)^2=0.$$

The roots of this equation are

$$\lambda=1-2\alpha^2+2\alpha^2\beta^2\pm 2\alpha\beta\sqrt{1-\alpha^2(2-\beta^2)}.$$

In figure 4, the curve $1=\alpha^2(2-\beta^2)$ is shown by a chain line. For $1<\alpha^2(2-\beta^2)$ on the right-hand side of the chain line, the two roots become complex conjugate and their magnitudes are both equal to $|1-2\alpha^2|$ which does not depend upon the parameter β. For $1>\alpha^2(2-\beta^2)$ on the left-hand side of the chain line in figure 4, it can be shown that the two roots are real and positive, and one of the roots (larger one) becomes greater than or equal to unity depending upon $\beta\geq|\alpha|$.

In conclusion, the method II is unconditionally unstable for $0<|U|<C$ (i.e., a subcritical or subsonic flow) and

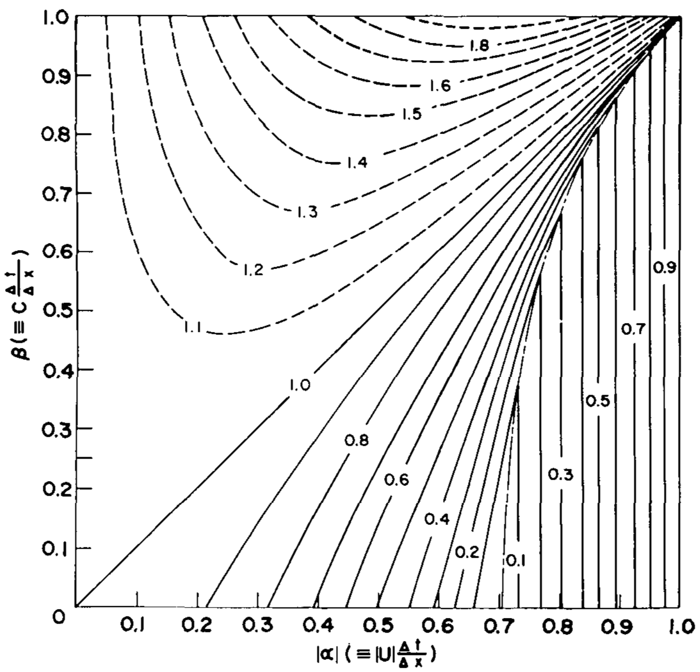

FIGURE 4.—Magnitude of the largest root ($\theta=\pi$) of equation (22), plotted against $|\alpha|$ and β (method II).

conditionally stable for $|U|\geq C\geq 0$ (i.e., a supercritical or supersonic flow) provided that $|U\Delta t/\Delta x|\leq 1$, the Lax-Wendroff condition. If U is zero, the stability condition is $C\Delta t/(2\Delta x)\leq 1$.

4. REMARKS

When one wants to solve a partial differential equation numerically, one must first write down a finite-difference form of the differential equation and then study the stability property of the difference equation before ever attempting to integrate the equation. It is customary to check the stability of the difference scheme in the von Neumann sense, that is the stability of the corresponding linearized system with constant coefficients. However, it is not always easy to obtain *analytically* the von Neumann condition for stability for the system in which many physical factors such as advection, gravity waves, dissipation, etc., are involved. It is tempting, therefore, to introduce approximations of various degrees in order to simplify the stability analysis. One of the common approximations is to check the stability of difference equations considering only one physical factor of the system at one time. Then one writes down a stability criterion inclusive of all the stability conditions obtained separately for every physical factor. By doing so, one simply hopes that the combined stability criterion is as good as the "complete" stability condition which would take into account all physical factors under consideration.

It was shown in this note that such a practice is a bad one through the demonstration of a counter example to this procedure. For such problems, it is recommended that the evaluation of eigenvalues of the amplification matrix should be performed analytically or numerically for various values of all the physical parameters involved in the system in order to determine the ranges of the physical parameters for which the eigenvalues are equal to or less than unity in magnitude for physically stable systems, and the eigenvalues do not exceed $1+O\ (\Delta t)$ for problems in which there is a mechanism permitting a growth of the true solution.

ACKNOWLEDGMENTS

A large part of this work was done while the author was visiting the AEC Computing and Applied Mathematics Center, Courant Institute of Mathematical Sciences, New York University in 1962–3. The author is indebted to Professors R. D. Richtmyer and E. Isaacson for suggestions and advice on this problem. Also the author has had the benefit of discussions with Drs. J. Gary and G. Fischer, National Center for Atmospheric Research, and Professor George W. Platzman, the University of Chicago.

REFERENCES

1. R. Courant, K. O. Friedrichs, and H. Lewy, "Über die partiellen Differenzengleichungen der mathematischen Physik," *Mathematische Annalen*, vol. 100, 1928, pp. 32–74.
2. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, II, Interscience Publishers, New York, 1962, 830 pp.
3. G. Fischer, "A Survey of Finite-Difference Approximations to the Primitive Equations," *Monthly Weather Review*, vol. 93, No. 1, Jan. 1965, 1–10 pp.
4. P. D. Lax and B. Wendroff, "Systems of Conservation Laws," *Communications on Pure and Applied Mathematics*, vol. 13, 1960, pp. 217–237.
5. R. D. Richtmyer, *Difference Methods for Initial-Value Problems*, Interscience Publishers, New York, 1957, 238 pp.
6. R. D. Richtmyer, "Proposed Fluid-Dynamics Research," New York University, Institute of Mathematical Sciences (S-600928), 1960, 56 pp.
7. R. D. Richtmyer, "A Survey of Difference Methods for Non-Steady Fluid Dynamics," National Center for Atmospheric Research, *Technical Notes* 63-2, 1963, 25 pp.
8. J. J. Stoker, *Water Waves*, Interscience Publishers, New York, 1957, 507 pp.

[*Received April 23, 1964; "Remarks" added November 3, 1964.*]